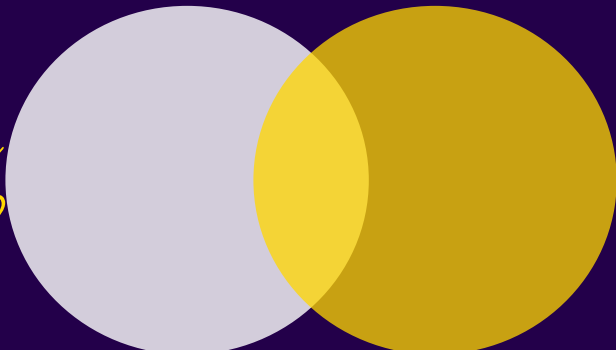
DISCRETE MATHE-MATICS *without* FORMULAS

Alexander Solovyov



DISCRETE MATHEMATICS WITHOUT FORMULAS

Lectures without a board

(For those who can't stand formulas, but have imagination)

Translated from the Russian and
Typeset in LaTeX by *Damitr Mazanav*
damitr@proton.me
Released on the Web in 2025.

TRANSLATOR'S NOTE

The original book was typeset in Word 2000. I have tried to maintain the original setting, but have made some stylistic changes. At times I have added clarificatory margin notes. And some illustrations. I hope, I have maintained the spirit in which the book was written, without formulae and without much formal mathematical symbolism.

Damitr Mazanav
September 2025

Contents

A Lyrical Digression		vii
1	Sets	1
2	Infinity Comes in Different Forms	11
3	Operations On Sets	19
4	Correspondences, Mappings, Relations	27
5	Special Relations	35
6	Mathematical Logic (Meaningless Logic)	43
7	Operations On Statements	49
8	Propositional Algebra	55

9 Predicate Logic 61

10 Axiomatic Theories 67

11 Theory Of Algorithms 75

12 Formal Grammars 85

13 Computational Complexity 91

The author expresses his gratitude
to Word (or Ward) – 2000,
which completely killed his desire to
write formulas.
Absolutely.

A Lyrical Digression

Formulas in mathematics were invented, strangely enough, to make mathematics easier. Schoolchildren still refuse to believe this.

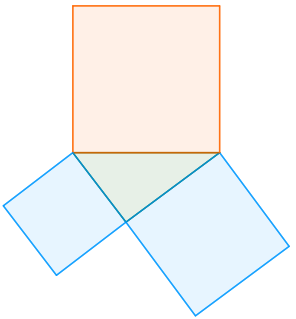
In ancient India, for example, they managed quite well without formulas: they took a monkey, a jug, a banana, a woman, and other tangible objects… And, as we would say today, built logical conclusions in a concrete manner. The Egyptians, instead of formulas (which had not yet been invented), carved beautiful reliefs in their deserts — a kind of formulaic drought. And the Arabs once practised mathematics without formulas as well—they did it in verse. They composed melodious poems about the square of a sum …

But then things took a turn for the worse.

The blame lies squarely with Pythagoras, who declared that "EVERYTHING IS NUMBER!"…It is hard to think of a

greater falsehood in mathematics, even today! Although, to be fair, there is nothing wrong with Pythagorean pants[1] — those were designed to last for centuries.

[1] This refers to the graphical representation of Pythagorean proposition depicting *Euclid's I.47* commonly termed as Pythagorean pants in Russian. – DM



Many people dislike mathematics today. The overwhelming majority of the general public. And they mostly dislike it because of formulas. Rightly so. Because back in school, they were warned that mathematics is formulas, just as poetry is rhyme. Which is why they do not engage with mathematics but, instead, spend their days sending rhymed greetings to their friends on the radio, trying to astonish the world with their artistic revelations:

> *"Wishing you joy, may troubles pass by",*
> "A hundred years without tears."

Brilliant, isn't it?

Imagine my surprise, then, when I read books on the foundations of mathematics and found… almost no formulas.

Formulas, of course, make life much easier for a mathematician, but this is something one must come to appreciate voluntarily, not through forced *formulaic torment* — *as if to say, endure it, and you'll grow to love it*…

All our humanitarian problems in mathematics stem not only from some crackpots' obsession with abstraction. There is an even more insidious word here: SEMANTICS, which, in plain human language, usually means MEANING. So, first of all, mathematics usually struggles with meaning by all means

available to it, including formulas… Of course, in the name of achieving *the* HIGHEST MEANING. As is always the case with us!

The more "advanced" mathematics is, the less SEMANTICS it contains and the more SYNTAX. Everyone studied syntax in school, and to this day, it brings little joy to anyone!… Perhaps T. Tolstaya is right when, in her novel Kys, she refers to "syntax" with a swear word. Perhaps T. Tolstaya was right when, in her novel *Kys*, she used "syntax" as a swear word.

But this is not yet a lecture — just a lyrical digression.

P.S.

Since the advent of the internet, a lot of very smart people have appeared online, often with advanced degrees in physics and mathematics. This is good, but sometimes it gets tiresome … because not everyone, despite their great intellect, retains the ability to judge "according to the laws of the genre."

So, a special request to them: do not suspect the author of attempting to write a textbook on *foundations of mathematics* or a *monograph*. And the main request — do not read what is written below at all.

It is not for you.

Go in peace, back to where you came from…

1 Sets

What a "set" is should be clear from the word itself, without any definition. Moreover, providing a definition for this fundamental mathematical concept is impossible. Don't even try.

Better spend your energy on a perpetual motion machine or something more concrete…

A set can be a collection of trees in a forest, a group of students at a university, or even a number of poor relatives in America who might send you an invitation… There is, of course, a very serious game called "AXIOMATIC SET THEORY." Only a select few can grasp its rules, and no one has found a practical application for it… But this is entertainment for a very closed circle of enthusiasts, given that the theory itself is also quite closed.

A set consists of elements — trees, students, poor rela-

tives… It does not matter whether we list the students alphabetically or by academic performance.

Duplicates, however, are not allowed, nor are students who lack distinguishing features. Even if you are Chinese students, you must differ from one another… There can even be sets consisting of numbers. But, as we agreed, we shy away from mathematics in general and from numbers in particular as the devil avoids holy water… So we can do without them. Or with them. Or we can use only those numbers that have been familiar to us since childhood…

However, we will not consider "a collection of thoughts in one's head" to be a set. Not because of their number, but because these thought-elements cannot be clearly separated in the general mess, neatly arranged on shelves, or properly labeled. A set of thoughts, if sorted onto shelves, simply would not fit in one's head due to the outdated design of the average brain.

By the way, since the Russian word for "set" (множество) suggests "many," and the idea of "many" (*many*) is subjective, to avoid disputes among Russian speakers, we will use "set" for any number of elements, just as English speakers do. Even for a single element. Even when a set has no elements at all — such a set is called *empty*! This, in particular, allows one to correctly tell their friends, from the perspective of set theory, a joke about "the set of needy veterans of the Battle of Kulikovo…"[2]

[2] The Battle of Kulikovo (1380) was a historic battle between the Russian principalities and the Golden Horde. – DM

Besides the concept of a set, there is only one other fundamental concept — and that's it. Everything else in this theory is derived. This second fundamental concept is MEMBERSHIP (or the "relation of membership"). That is, an element belongs to a set. There is no need to define this explicitly, as the word belongs in everyday speech can be replaced with many synonyms depending on the context, such as:

§ That birch tree "is located" in this forest,
§ Sidorov "is listed" as a student,
§ Mr. X "is included" in your poor American relatives.

Note: To avoid synonyms that might confuse us, we could introduce a special small symbol resembling the Greek letter epsilon. But we won't do that, because from this symbol to full-fledged formulas is just a short step…

Important Warning: Questions like "Does student Sidorov belong to the set of balding people?" take us away from set theory, so we will simply ignore such questions, rightly considering that classical set theory does not deal with baldness — since there are no objective criteria for baldness. Therefore, the issue of membership or non-membership can be resolved informally, perhaps for a fee.

The solution here is quite simple: first, establish a clear definition of baldness somewhere else in an official capacity, and then bring in set theory.

In other words, we always assume that we clearly know what belongs to a given set and what does not! Anything else

is considered non-existent altogether!!!

Furthermore, if we want to say that all birch trees (a birch tree, unlike a balding person, is still a birch tree even in Africa) found in a given forest also belong to the entire forest wealth of our country, and that all students enrolled in a university are also counted as students of Russia, we use the terms *subset* or *included* to shorten such statements.

Here too, there may be obvious synonyms. But to avoid confusion and not to mix them up with *belongs to*, we must remember one simple rule: *belongs to* applies when an "ELEMENT belongs to a SET", whereas *included* applies when a "SET is included in another SET". That is why the alternative term for "included" is *subset*, meaning a certain part of a set.

The set of university students is *included* in the set of students of the country. That is, the set of university students is a *subset* of the set of students of the country.

Anyone who did not twist their tongue while reading this will understand that the set of students of the country is *included* in the worldwide set of students.

We can continue this chain of inclusions, extending it to the galaxy. But then it follows that the set of university students is a subset of the set of students in the galaxy.

This property of chains is simple and rigorously(!) proven directly based on how we defined the inclusion relation.

The inclusion relation has several interesting properties —

not invented by us. They can be discovered by any researcher who "plays around" with this relation.

For example, we can say that the set of students in group UH-001 is included in the set of university students since such a group is officially listed at the university. The fact that all students have been expelled from the group does not matter mathematically. Since there is NOT a single student belonging to this group who is not also listed at the university, such reasoning applies perfectly to all empty sets. This allows us to reach the general conclusion that the empty set is included in any set, including itself.

Appreciate the mathematical beauty of the following phrase:

> *"Any element belonging to a set that contains no elements also belongs to any other set that contains no elements."*

A slightly less elegant phrase is:

> *"Any set is a proper subset of itself."*

Or, the same idea but phrased more harshly:

> *"Any set is included in itself."*

Indeed, group UH-002 (which may very well have students) is included in group UH-002 because all the students listed in it still remain listed in it, even if its name, UH-002, is mentioned multiple times.

From this last example, we can draw an important conclusion. If two sets (which may initially appear different, like

the set of officials and the set of servants of the people) are included in each other, then these sets are *equal* — that is, they consist of the same elements.

Put slightly differently:

> *"If two sets are subsets of each other, then they consist of exactly the same elements."*

How could it be otherwise?!…

Of course, there are dissident mathematicians who do not accept this. But that is more a matter of faith… of another mathematical denomination…

Now, we must admit that mathematicians are akin to that category of afflicted individuals known as "truth-seekers". Generally speaking, seekers of (social) truth are right. But their correctness is either meaningless or unrealistic, and most importantly, it is needed by no one except themselves… Similarly, in set theory, one can often find truths that, to outsiders, may seem, to put it mildly, strange or even harmful.

For example, student Khvedorov *cannot* be a subset of the university's students because he himself is not a set but an *element*. As an element, he can only be an *element* of the set of university students. However, group UH-003, being a set of students, is a legitimate *subset* of the set of university students. But this group UH-003 consists of only *one* non-expelled student—Khvedorov himself!

So, it turns out that Khvedorov *cannot* be a subset, but the

group consisting solely of him *can* be.

On the other hand, if the university rector suddenly decides to treat the university as a set of student groups, then group UH-003 will become an *element* of the set of university student groups.

There is nothing confusing about this — provided one understands that the "set of university students" and the "set of university student groups" are two entirely different sets.

However, bureaucratic twists and turns do not surprise us—we see far worse every day…

Still, set theory does have something that can astonish even us. These are the so-called *paradoxes of set theory* — one of the great shocks of the first year of the last century, at least for a small group of people.

Let us illustrate this with the famous example of the barber.

A ruler (much like Peter the Great) issued a decree:

> the *only* barber in his kingdom must shave *all and only* those who do *not* shave themselves. Any violation of this order would be punished by death.

So, the barber set about shaving all the unshaven men. Eventually, he found himself faced with a dilemma — his own beard had grown out…

To shave or not to shave, that is the question… A man trying to shave himself with a blunt razor. Lithograph by William Green after M.W. Fry, ca. 1820. Public Domain Wellcome Collection

He took up a razor. But if he *shaves himself*, that means he is someone who shaves himself — whom he is *forbidden* to shave.

So, he put the razor down. But then he realised — he *does not* shave himself, which means he *must* shave himself!

And then what? Nothing good! The poor man is doomed

to execution, no matter what he does.

From the perspective of set theory, the barber's predicament arises because he cannot decide on a FUNDAMENTAL QUESTION OF MEMBERSHIP:

> *should he include himself in the set of those who do not shave themselves?*

In other words, at the very foundation of set theory — intended to be the bedrock of *all* mathematics — the fundamental concept of MEMBERSHIP can lead to contradictions so severe that they make it impossible to define certain sets at all!

Mathematicians arrived at the only reasonable solution: they simply *agreed* not to construct, within set theory, sets that *cannot* be constructed!

Thus, set theory operates on *all* sets — except those that cannot exist. All such non-existent sets, taken together, are called the UNIVERSE.

2 *Infinity Comes in Different Forms*

The most interesting aspect of set theory is that it considers not only finite sets — those containing a finite number of elements — but also infinite sets, for which even the concept of number becomes meaningless. In other words, set theory can deal not only with the set of students in a group or the set of birch trees in a forest but also with the set of points on a line and the set of stars in the sky…

The founder of set theory, Georg Cantor, suffered greatly because of infinity — so much so that psychiatrists had to intervene. Yet, long before him, mathematicians had been working with infinity. Consider the infinitely large set of points on a line or, conversely, the infinitesimal quantities in higher mathematics…

The problem is that no living person has ever seen, heard, or touched infinity! That is why, before Cantor, mathemati-

cians recognised and used only what is called POTENTIAL INFINITY. The most classic example is the concept of an infinitely large number in higher mathematics. An infinitely large number is one that is greater than any given number. If someone does not understand what that means, they are asked to name the largest number in the world!… An educated person might say a billion billion. But then they are told that an infinitely large number is greater than this — "even more than by another billion billion."

That means we always have a number potentially (!) greater than any number an erudite might come up with…

Cantor, however, introduced ACTUAL INFINITY into mathematics. Until then, only poets had been allowed such a luxury — poets, who, as we know, are never held to strict account…

> *"The stars are countless, the abyss has no bottom."*

Poets dislike taking things bit by bit, drop by drop… They prefer everything at once!

> *"Behold, the entire abyss of your downfall!…*
> *I give you all the stars — such a trifling gift is nothing for you, my priceless one and only!"*

According to Cantor, infinity exists as a whole, all at once. And since infinite sets exist in their entirety, they can be manipulated mathematically. They can even be compared in terms of "greater" and "smaller".

Thus, Cantor began asking himself "poetic" questions and

seeking mathematical answers to them. One of the key questions was:

> "Is *infinitely many* always the *same infinitely many*? Or *can some infinities be larger than others?*"

Which is greater: the number of stars in the sky or the number of points on a line?…

Cantor proved a great theorem, which shows that infinities can differ in size. Since terms like "number" and "quantity" are inadequate in this context, he introduced the concept of *cardinality*. Cardinality is what remains when we disregard the nature of a set's elements and the order in which they are arranged. He defined this concept precisely, even though the definition may seem strange at first glance. Upon a second look, it usually no longer seems so. If we were to stop distinguishing students and ignore their order, only their cardinality would remain.

Alas, it is impossible to give examples of sets with infinite cardinality using birch trees or students, and the stars are far away and only visible at night. Therefore, for clarity, let's turn to the numbers that are close at hand.

When we count objects, we use whole (positive) numbers: 1, 2, 3… These are called *natural numbers*. Strange Americans like to start this sequence with zero (and have infected, for example, all of computing with this habit). They are unbothered by the fact that a "3-block" is actually in the fourth row… But never mind! Adding or removing zero does not change anything.

The important point is that we know we have enough numbers to count anything. We also know that this set is infinite. Cantor called this set COUNTABLE and its cardinality — the cardinality of a countable set.

Cantor took this cardinality as a standard and began comparing it with the cardinalities of other sets.

First, he established that this cardinality is greater than that of any finite set (students, birch trees, etc.).

Second — and this is particularly interesting — he proved that many infinite sets have the same cardinality (the same "number" of elements) as the set of natural numbers. One of the most astonishing examples is that the set of all positive integers has the same number of elements as the set of all positive even integers! In other words, they are *equinumerous*.

Indeed, let us write them out, one below the other:

Natural Numbers	1	2	3	4	…
Even Numbers	2	4	6	8	…

It is clear that both sequences contain the same number of elements because each number in the first sequence ALWAYS corresponds exactly to one number in the second sequence. Therefore, the second sequence cannot be exhausted before the first—and vice versa!

Thus, these sets have the same cardinality!

Therefore, here the *part is equal to the whole*!

Since this has been rigorously proven, the last desperate argument — "this does not happen in real life" — can be countered even more strictly:

> *"You simply haven't seen life! More precisely, you have never seen infinity in real life! And you never will!"*

Over a long and complex life, a person may encounter a locomotive, but infinity — never! Even in the dark.

Therefore, what can and cannot exist in the world of infinities is not for us to judge based solely on everyday experience!

From the infinite set of stars (which also has countable cardinality), we see only a limited, finite subset. On a drawn line segment, which contains an infinite number of points, we see only a finite number of graphite grains that form the segment. By the way, we see all this and much more with the retina of the eye, which contains a finite number of rods and cones. With a finite number of rods and cones in your eye, you will never see anything infinite!

Thus, infinities around us exist in a "parallel world" governed by their own laws, which set theory helps to study.

We have already mentioned the "second point," but there is also a "third" one — and this third is the most important: Cantor's great theorem, which has already been mentioned.

The essence of this theorem is that if we construct the set of all subsets of a given set, we always obtain a set LARGER than the original one.

For example, let us take a set with two elements: ONE, TWO (and that's all). The subsets of this set will be *four* sets(!):

1. ONE, TWO – (any set is a subset of itself)
2. ONE
3. TWO
4. The empty set – (i.e., "that's all").

Another example: A & B (sat on a pipe). The subsets of this three-element set will be eight sets:

1. A, &, B
2. A, &
3. A, B
4. &, B
5. A
6. &
7. B
8. The empty set

For a four-element set, there would be 16 subsets. This pattern can continue infinitely, following the sequence of powers of 2.

Now, Cantor proved that if we take an infinite set of countable cardinality, for example, the set of all positive integers, and construct (purely conceptually) a set whose elements are all subsets of this set, we obtain a cardinality *greater* than countable infinity.

In principle, there is no way to enumerate (even in infinity) such a set. It will always contain more elements. This new,

larger cardinality is called the *cardinality of the continuum*.

And again, a paradox of everyday reasoning. The *cardinality of the continuum* is, for example, the cardinality of the set of points on a line or the set of real numbers, which are essentially the same. Moreover, any segment of the number line, even a tiny segment like the interval from 0 to 1, has the cardinality of the continuum, meaning it contains more numbers than any countable set. Since this interval has the cardinality of the continuum, just like the entire (infinite) number line and, naturally, any of its segments, one can say that the segment from 0 to 1 contains exactly as many points as the segment from Earth to Jupiter.

Here too, a *part equals the whole*, as long as both the part and the whole have the cardinality of the continuum. And they all equally exceed the number of stars in the sky or the number of all possible algorithms.

For infinities, there is a very simple arithmetic that logically follows from our previous discussions. Adding two countable cardinalities results in a countable cardinality, and for continua, it results in the cardinality of the continuum. Subtracting a countable cardinality from the continuum still leaves the cardinality of the continuum. However, subtracting one continuum from another, or one countable set from another, can yield different results in each particular case. One could easily come up with their own illustrations here.

However, things are not so simple. Infinity remains one of

the key concepts in philosophy. And mathematics only fuels the fire by revealing new facets of this problem. Especially when considering not just infinite sets but also infinite *ordered* sets. Those interested may read books on some very elegant topics with equally beautiful and romantic names: *Cardinals & Ordinals*.

3 Operations On Sets

The term *operations on sets* is used not because these operations are physically *above* sets, but simply because that is the accepted terminology. If a hairdresser performs operations *on* your hair, it does not mean the result of their work will be positioned above your hairstyle. (But beware of a surgeon performing operations *on* patients.)

There are only three fundamental operations. This is fewer than in school arithmetic. In fact, even this set of operations is somewhat redundant. These operations are called UNION, INTERSECTION, and COMPLEMENT. They resemble, to some extent, the school arithmetic operations of addition, multiplication, and sign change. However, this analogy is rough and can be misleading — after all, it is just an analogy.

Let's start with a historical anecdote.

Axel Ivanovich Berg — an admiral, academician, and a

man with a fiery temperament — was one of the first advocates of cybernetics in the USSR, at a time when it was still officially considered a "corrupt handmaiden of capitalism." At that time, discrete mathematics was not taught in technical universities due to its supposed lack of practical value, though cybernetics was beginning to make timid use of it.

During a conversation with a "science journalist," who argued that set theory was not only unnecessary but also incomprehensible to the average Soviet engineer, Berg abruptly interrupted the discussion and ordered his driver to take them to the nearest kindergarten.

At the kindergarten, the children were playing in a large sandbox — one of the few entertainment options available in post-war kindergartens. Berg drew two large, partially overlapping circles in the sand, similar to how wedding rings are depicted on greeting cards and cars. (For those unfamiliar with weddings, a similar overlap is seen in the Olympic rings.)

He then said: "Let everyone who likes semolina porridge stand in the left circle, and everyone who likes plum jelly stand in the right one!" The children, eager to eat (as it was a time of post-war hunger), did not hesitate—every single one of them ran into the drawn circles.

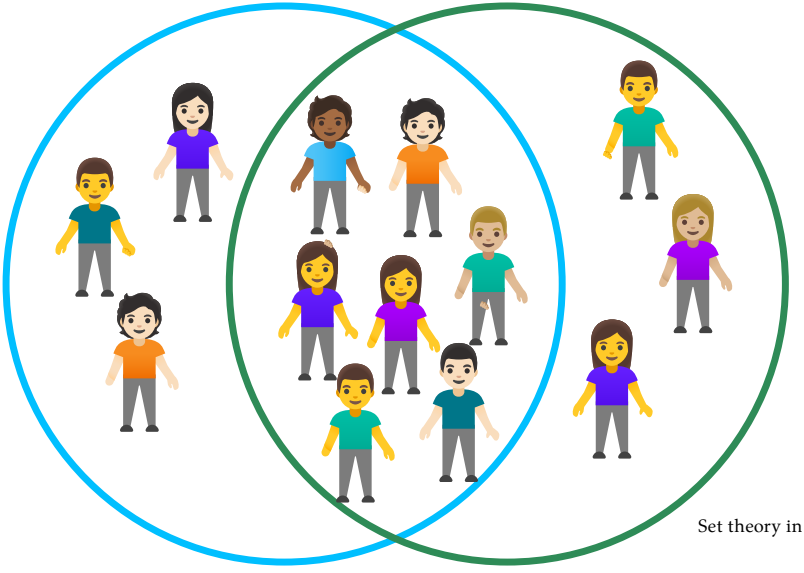
The union of all these little sweet tooths represented the *union operation* in set theory.

However, since almost all the children stood in the overlapping area of the circles — indicating their love for both

porridge and jelly — they unintentionally demonstrated the *intersection operation* of two sets.

"Well then! I don't know about engineers, but children certainly understand the meaning of set operations!" Berg declared.

Incidentally, in this case, the entire sandbox played the role of the *universal set*.



Set theory in the kindergarten.

What Berg had drawn in the sand is now known as EULER-VENN DIAGRAMS. Everything outside the circles in the sand represented the *complement* of the corresponding set — i.e., the elements of the universal set that did not belong to the group of food lovers. (This included Berg and the journalist.)

If we carefully examine student group UH-004, the union of the set of top students and the set of athletes will form a new set called "the pride of group UH-004."

The fundamental difference between set union and regular addition in school arithmetic is not just that students are not numbers and we do not count them(!), but also that students who are both top students and athletes are *counted only once*. Thus, it may well happen that there are four top students and twenty athletes, yet their union — "the pride of group UH-004" — contains only twenty-two students.

Clearly, the *intersection* of these sets will consist of two students who are both top students and athletes. They are most likely girls — and probably beautiful ones at that — but beauty was not used as a criterion for defining these sets…

When mathematicians have objects at their disposal — and here we have complete freedom to choose any objects — and when they define operations (having already introduced the main three), they begin to speak of ALGEBRA.

Set algebra is as different from school algebra as the sky is from the earth, though there are some analogies. In set algebra, the same law names appear: COMMUTATIVE, ASSOCIATIVE,

and DISTRIBUTIVE (also known as the *permutation*, *combination*, and *distribution* laws). The first two are as similar as two drops of water falling from the sky to the ground. The distributive law, however, has both an analogy in school algebra, where the product of a sum is the sum of the products, and a unique version. In set theory, to put it briefly, the intersection with the union equals the union of intersections, and (!) the union with the intersection equals the intersection of unions. The second has no analogy in school algebra: "The sum with the product is *not equal* to the product of sums."

Let us illustrate these laws.

Commutative Law: The union (intersection) of top students and athletes is equal to the union (intersection) of athletes and top students.

Associative Law: Changing the order of the union (intersection) of athletes, top students, and beauties does not change the result.

Distributive Law: (only the exotic version): The union of beauties with the intersection of athletes and top students is equal to the set where the unions of beauties and athletes intersect with the union of beauties and top students. (In symbolic notation, this would be much shorter and clearer, but we've sworn off formulas).

The *absorption law* is somewhat difficult to grasp by ear, but it allows for the simplification of set-theoretic constructions in certain cases. The intersection of top students with the union

of top students and athletes results in the set of top students. Another version states that the union of top students with the intersection of top students and athletes also results in the set of top students. However, if one thinks about it carefully and gestures a bit, the correctness of these results becomes obvious.

There is another law whose name seems to amuse students, perhaps because it reminds them of something. Yet this law can confidently be considered one of the most important. It is the *idempotent law*, which states that the union or intersection of a set of athletes with itself results in the set of athletes.

The *De Morgan's laws*, which sound very French, state that the complement of the union of top students and athletes is equal to the intersection of the complement of the set of athletes with the complement of the set of top students. The second version states that the complement of the intersection of top students and athletes is equal to the union of the complement of the set of athletes with the complement of the set of top students. The universal set (for taking complements) can be considered as the group of students, the university, or even the whole world — it does not matter. Take real athletes and top students and verify the validity of the law for yourself.

The law of *double complement* is very simple. The complement of the complement of the set of athletes is the set of athletes itself. For those who have successfully navigated through our verbal casuistry, we can state an immediate consequence of this law:

> *"the complement of the complement of the complement of the set of athletes is simply the complement of the set of athletes."*

The most exotic laws are the *law of contradiction* and the *law of excluded middle*.

Law of contradiction: The law of contradiction states that the intersection of the set of athletes with the complement of the set of athletes is empty. Indeed, since the complement of the set of athletes includes all other students who are not athletes, there can be no common elements in this intersection.

Law of excluded middle: The law of the excluded middle states that the union of the set of athletes with the complement of the set of athletes coincides with the given universal set. Indeed, since the complement of the set of athletes includes all other non-athlete students from the universal set, this union forms the entire universe.

It is unfortunate that not all mathematicians agree with these laws. Even more unfortunate is the fact that they have very strong reasons for their disagreement — just as strong as those held by the supporters of these laws.

The dissenters call themselves CONSTRUCTIVISTS or INTUITIONISTS. The ones who accept these laws had no choice but to call themselves CLASSICISTS — something with which the dissenters do not agree.

4 Correspondences, Mappings, Relations

Algebra does not nearly exhaust everything that can be done with sets…

In mathematics, as in life, various objects may correspond to something or not. They may be in certain relationships with one another or, conversely, not be. The foundation for formalising — or, if you will, "mathematising" — such connections also relies on sets.

That is, different *correspondences* and *relations* can be established between sets. Moreover, and perhaps more importantly (serious mathematicians might even say "first and foremost"), sets can often be *mapped* onto each other and even onto themselves.

A person may correspond to a profession, a salary to a position, a punishment to a crime, and a grade to knowledge.

Looking at numerous examples around us, we notice that to define a specific correspondence, we must determine two sets: the *domain* (set of inputs) and the *codomain* (set of outputs). We also need to define "correspondence pairs".

For example, the domain is group UH-005, which is taking an exam. The codomain is the set of grades: excellent, good, satisfactory, unsatisfactory. The set of pairs could be: Ivanov – excellent, Petrov – good, Sidorov – excellent. But Khederov did not show up. And there you have a defined correspondence.

Correspondences have properties.

1. In this case, the correspondence is *not everywhere defined* because there is no pair for Khederov in this correspondence. Even if we had written "Khederov – absent" in the record book, it still would not be part of the correspondence since "absent" is not in the set of permissible values. If the dean's office had promptly removed Khederov from the list as an expelled student, the correspondence would have become *everywhere defined*.
2. The correspondence is *functional* because each student corresponds to no more than one grade. Such a correspondence is commonly called a *function*. However, due to Khederov, this is not an everywhere defined function. There is no difference from a school-level function, except for the fundamental point that, here, both arguments and values can be not only numbers but also any objects. Interestingly, not all mathematicians like this definition of a function, although it is absolutely rigor-

ous. This is due to a certain rivalry between set theory and other branches of mathematics. If students could receive multiple grades for a single exam, the correspondence would be non-functional, meaning it would not be a function. (It would be a "multivalued [nondeterministic] function," but that belongs to a different branch of mathematics.) In real life, such cases do not occur.

3. This correspondence is *not injective* because more than one student received an "excellent" grade. If Sidorov, due to some unfortunate fate, had received a "satisfactory" or "unsatisfactory" grade instead of "excellent," the correspondence would have been injective. The awarding of Olympic medals for winning a 100-metre race would be an example of an injective correspondence.
4. The correspondence is *not surjective* because not all possible grades were used in the exam. In real exams, the full range of grades is typically assigned, making the correspondence *surjective in practice*. Naturally, ticket purchases for a Vitas concert are currently surjective.
5. A correspondence that is *everywhere defined*, *functional*, *injective*, and *surjective* is called *bijective*. It is also referred to as *one-to-one and onto*, though that sounds less elegant. The most compelling example of a bijective correspondence is the number of heads on shoulders. Consider the set of heads and the set of shoulders, and verify all four properties. No criminal cases allowed!

The classification of correspondences as a separate category was proposed by Europeans, more precisely, the French, and

even more precisely, Nicolas Bourbaki (a collective pseudonym for a group of intellectual mathematicians, akin to the Russian satirical writer Kozma Prutkov[3]). The American school, on the other hand, considers correspondences a special case of relations. Since we are treating relations separately, as it makes things more structured, it is now time to talk about relations.

[3] Kozma Prutkov is a fictional character created in the mid-19th century by a group of Russian writers, including Aleksey Tolstoy and the Zhemchuzhnikov brothers (Alexei, Vladimir, and Alexander). Prutkov was conceived as a satirical persona, embodying the archetype of a pompous, self-important bureaucrat and pseudo-intellectual. He is known for his absurdly profound aphorisms, poems, and fables, which mock the pretentiousness and incompetence often found in government officials and literary figures of the time. – DM

In mathematics, as in life, different objects may or may not have a relationship with other objects.

Family relationships, friendships, diplomatic relations, and equal relations.

Looking at numerous examples around us, we notice that relations differ from correspondences in that they are defined within a single set. It would be meaningless to talk about relations between students and grades or about diplomatic, familial, or any other kind of relationship between job positions and salaries. To define a specific relation, we need to determine the set and the pairs for which this relation holds.

For example, within the set of people, relations could include "being a brother," "studying in the same group," or "being taller."

Relations, due to their nature, are characterised by a different set of properties than correspondences.

1. **Reflexivity.** This is when a relation is directed at itself. The inclusion relation was previously discussed. Since any set includes itself, the inclusion relation has the prop-

erty of reflexivity. If folk wisdom is to be believed, then the relation of "saving" among drowning people is also reflexive.

2. **Antireflexivity.** This is when a relation is never applicable to the object itself. For example, "perpendicularity" among lines. A line cannot be perpendicular to itself.
3. **Symmetry.** If Ivanov "studies in the same group" as Petrov, then the reverse is also true. If line A is "perpendicular" to line B, then the reverse is also true.
4. **Antisymmetry.** If a thousand roubles can be "exchanged" for hundreds, the reverse is beyond even a magician's ability. A grim but precise example: "mourning someone"…
5. **Completeness.** This is the most complex property because, unlike all the others, it is primarily "directed" at the set itself. A relation is *complete* if it holds for any pair of different elements in the set in at least "one direction." For example, the relation "greater than" for the set of real numbers is complete because, for any two distinct real numbers, one must be greater than the other. However, if we extend the set to include complex numbers, the completeness property disappears. If even one of the numbers being compared is complex, the comparison in terms of "greater than" or "less than" loses meaning.
6. **Transitivity.** If Ivanov "studies in the same group" as Petrov, and Petrov studies with Sidorov, then Ivanov "studies in the same group" as Sidorov. The inclusion relation is also *transitive*. If a group is "included" in the

set of university students, and this set is "included" in the set of students in the country, then the set of group students is "included" in the set of students in the country. This chain of inclusions could be extended to the galaxy. And here again, we encounter a hidden pitfall of logical casuistry!

If a student group is considered an element of a university — a set consisting of groups — and a university is considered an element of higher education — a set consisting of universities — then the group is not an element of higher education (since its elements are universities). That is, the relation of "belonging to" is not transitive. "The vassal of my vassal is…"

Returning to functional correspondence (that is, to a function): if this correspondence is also everywhere-defined, it is called a MAPPING.

If we map the set of students in a group onto the set of surnames in the group, this is likely a *mapping* of the set of students *onto* the set of surnames. That is, a surjective correspondence. However, if we map the set of students in the group to the set of surnames of university students, it is said that there is a *mapping* of the set of students *into* the set of surnames. That is, the codomain will contain "unused surnames."

We have now approached one of the most fundamental (perhaps, for that reason, not the most euphonic) concepts in set theory and mathematics in general — we have arrived at

HOMOMORPHISM.

Example. Let us map the set of points on a section of the Earth's surface to the set of points on a map. For now, we will set aside the fact that certain sets of points on the Earth's surface may be mapped to a single point on the map — non-injectivity is common in such cases. What matters to us is that the higher the points on the Earth's surface are above sea level, the more brown the corresponding points on the map become.

Thus, we are not merely considering sets of elements. In the first case, there exists a relation "higher" between elements of the set, and in the second, a relation "browner." Where a point is higher in the first set, it is browner in the second. "Higher" and "browner" are relations defined on their respective sets.

The mapping of the Earth's surface *onto* the map does not simply associate elements of one set with elements of another. In addition, if a "higher" relation exists between two elements of the first set, then a "browner" relation holds between their images in the second set. Naturally, if points on the Earth's surface lie at the same altitude, they will be mapped to points of equal brownness on the map.

Such a mapping is called *homomorphic*, or we say that a *homomorphism* exists between these sets.

Returning to the fact that the word is not particularly euphonic—and, by American standards, rather cumbersome — recently, the shorter, truncated term MORPHISM has become

more common.

Morphisms play an exceptional role in mathematics. Since mathematics is often equated — quite justifiably — with mathematical modeling, let us conclude with an aphorism from an insightful philosophical book:

> A GOOD MODEL IS ALWAYS HOMOMORPHIC.

An aphorism at the end of a lecture provokes reflection — which is precisely the goal…

5 Special Relations

Each specific relation possesses a combination of properties. It is useful to study groups of relations that share the same set of properties.

Foremost among these are EQUIVALENCE relations. These are relations that simultaneously exhibit *reflexivity*, *symmetry*, and *transitivity*. The relation of EQUALITY between numbers is the simplest example of equivalence. Another example is "studying in the same student group."

Interestingly, every object is equivalent to itself, if only because even the most extraordinary object — completely unlike any other — still satisfies reflexivity, symmetry, and transitivity with respect to itself. Typically, however, objects are not so unique, and what we observe are sets (often called CLASSES) of *mutually equivalent* objects.

The most important property of an equivalence relation —

one defined precisely by the three properties mentioned above — can be illustrated with an example.

Example. Consider an initial state of chaos, that is, the entire set of university students wandering through hallways, sitting in the cafeteria, attending lectures, or, better yet, staying at home (or who knows where). The relation "studying in the same group" *partitions* this set into subsets — student groups. Every student belongs to exactly one group and cannot belong to two at the same time. (In real life, exceptions to these obvious properties may exist, but we assume we are dealing only with normal students.)

As a laboratory experiment on partitioning, it is recommended to break a plate — preferably one made of fine Chinese porcelain. Then contemplate the fragments, each of which will be an *equivalence class* for the porcelain pieces with respect to the relation "belonging to the same fragment…"

This is certainly a better option than breaking up groups, especially considering that orthodox algebraists do not understand "group" as a mere collection of students, but rather as something fundamentally mathematical. However, this is already the beginning of another romantic story — one about the young French genius and, alas, duelist Évariste Galois.

ORDER RELATIONS also play a significant role in mathematics. These are relations that possess the properties of *transitivity* and *antisymmetry*. Violating either of these properties disrupts order not only in mathematics but also in common

sense.

Examples.

§ "Being greater than" in the set of numbers.
§ "Coming after" in a queue.
§ "Holding a higher rank" in the military.

Additionally, if order relations possess the property of *completeness*, they are called PERFECT ORDERS. For example, the "greater than" relation on the set of real numbers.

If an order relation is also *reflexive*, it is called a NON-STRICT (PARTIAL) order. An example is "being at least as tall as." The previous three examples, on the other hand, are STRICT (or LINEAR) orders, as they possess *anti-reflexivity*.

It is reassuring that set-theoretic order relations generally align with everyday notions of order — but not always. A famous counterexample is the relation "being a brother." At first glance, it resembles an order relation:

Ivan is a brother of Marya, but Marya is not a brother of Peter — this seems to satisfy anti-reflexivity. If Ivan is a brother of Peter and Peter is a brother of Marya, then Ivan is a brother of Marya – this seems to satisfy transitivity. However, if Ivan is a brother of Peter, then Peter is also a brother of Ivan, meaning anti-symmetry fails.

Worse still, if Ivan is a brother of Peter and Peter is a brother of Ivan, then transitivity would imply that Ivan is a brother

of himself, which is absurd. To avoid this contradiction, the relation "being a brother" is classified as *non-transitive*.

Even more intriguing are other relations that closely resemble order relations. For example, "being slightly taller than."

Liberation, M.C. Escher, c. 1955.

This relation is *anti-symmetric* but not *transitive*:

Ivan is slightly taller than Peter, Peter is slightly taller than Yegor, Yet Ivan is significantly taller than Yegor.

Relations that *resemble* order relations but lack *transitivity* are called TOLERANCE RELATIONS. A classic illustration of tolerance relations can be found in Escher's famous artworks, where, for example, one shape (birds, animals, geometric shapes) gradually (smoothly) transforms (metamorphosises) into another (birds, animals, geometric shapes).

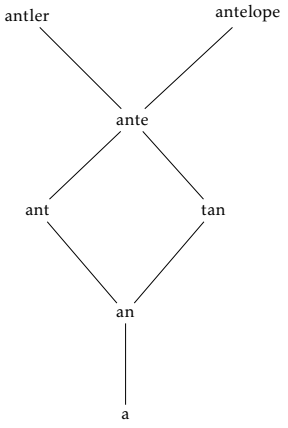
Partial order relations, meaning reflexive, anti-symmetric, and transitive relations that satisfy additional properties, are studied within a branch of mathematics with the exotic name LATTICE THEORY. This name can be intimidating, so in our country, the term "lattice" was initially translated as "structure." However, as the term "structure" became more widely used in mathematics, it had to be given its proper Russian equivalent, and lattices remained lattices in our language as well.

It is likely that the name "lattice" arose due to the use of HASSE DIAGRAMS, which may resemble ornate window

lattices. However, we have agreed to avoid formulas — and even more so, drawings. Unlike formulas, people tend to like diagrams, but drawing them in Word is even more tedious than writing formulas. Therefore, we will attempt, as much as possible, to compensate for their absence with verbal clarity.

Let us begin with examples of lattices.

Example. Consider the words: “a”, “an”, “ant”, “tan”“ante”, ”antler”, “antelope” and arrange them based on their inclusion within one another (keeping in mind that each word is included in itself). This will be our FIRST LATTICE.



Our first lattice.

It can be verified that all the properties of partial order hold in this example. As for additional properties, we will discuss them later.

The numbers: 1, 2, 3, 4, 6, 9, 12, 18, and 36, with the relation of divisibility, also form a lattice.

Ordinary real numbers with the relation “greater than or equal to” give one of the most common lattices. Although, for us, it is less exotic. One might say, as simple as a log…

The set of all subsets of a given set, with the *inclusion relation*, also forms a lattice, possessing several remarkable properties.

To define a lattice, let us agree to call an element the *greatest* (or *least*) if it is greater (or less) than any other element in a PARTIALLY ORDERED SET (POS). Mathematicians sometimes exhibit pedantic precision to the point of tediousness, and at

other times, an inexplicable looseness. Strictly speaking, it would be more precise — contrary to established tradition — to use the phrase "greater than or equal to" when referring to a POS that has the reflexivity property, calling it the *greatest or equal element*, and so on. However, we will also use the shorter term "greater", implicitly including these longer and more precise expressions.

The greatest element, if it exists, is *unique*. That is what makes it the greatest. The same applies to the least element.

An element of a POS is called *maximal* (or *minimal*) if there is no element in the set that is greater (or smaller) than it. At first glance, this definition seems to repeat the previous one. However, a POS may have *multiple maximal elements*.

Consider the modern world, ordered by levels of authority. All heads of state are maximal, but each within their own country, as there is no one above them. However, no head of state is greater than another. There is no single ruler of the entire planet, and even the UN Secretary-General does not serve that role. If, as in science fiction, a supreme leader of Earth existed, they would be both maximal and greatest.

Now, let us take a group of students in a class and impose a partial order on them — not by "any available means," but by the relation "performs better (or equally well) in studies." Let us assume that, for any two students, we can determine which one performs better. From this set, consider a subgroup UH-005 and identify the students in the class who perform

better than all members of UH-005. In other words, we find the *greatest* students relative to this group. There may be several such students unless the *greatest* student of the group is also the greatest in the *entire class*. The set of all such greatest elements is called the set of UPPER BOUNDS. Similarly, by considering the least-performing students, we obtain the set of LOWER BOUNDS.

Now, within the set of upper bounds (or lower bounds), we find the least (or greatest) element [this is not a typo!]. In this example, these will be the best and worst students of group UH-005. These elements are called the SUPREMUM and INFIMUM, or simply SUP and INF.

For the set of numbers {1, 2, 3, 4, 6, 9, 12, 18, 36} with the divisibility relation, consider the subset {3, 6, 9}.

§ The upper bound set is {18, 36}.
§ The lower bound set is {3, 1}.
§ The supremum is 18.
§ The infimum is 3.

A LATTICE is a partially ordered set (POS) in which every non-empty subset has both a *supremum* (least upper bound) and an *infimum* (greatest lower bound).

Lattices formed by the set of all subsets of a given finite set, ordered by inclusion, are called BOOLEAN LATTICES. For those familiar with Boolean algebra, the traditional Boolean algebra is a lattice consisting of two elements: "TRUE" and "FALSE", ordered by "more true."

A lattice can also be defined *algebraically*. If a POS satisfies the commutative, associative, absorption, and idempotent laws, then it is a lattice.

Furthermore, if the distributive law holds, the lattice is called a *distributive lattice*.

Believe it or not, lattices have been used to solve several important problems, including those in theoretical computer science.

6 *Mathematical Logic (Meaningless Logic)*

Usually, real mathematicians are not well adapted to life. Just look at them if you get the chance — they seem to be lost in their own world. One might assume that logicians should be an exception, at least because their actions ought to be the most logical. But that is far from the truth! Quite the opposite, in fact. Logic strictly defines its own "rules of the game" and follows them with ruthless precision, bordering on absurdity. However, its connection to the "logic of life" is no greater than the similarity between a game of chess and the Battle of Borodino. Yet, there is still something in common — when tensions on both sides rise to a boiling point, and things come down to hand-to-hand combat!

Incidentally, a classic comedic technique involves fanatically adhering to some formally logical but, from a real-world perspective, utterly absurd course of action. Everyone, especially children, can see the sheer idiocy of what is happening,

yet the clown persists, acting *logically*! And nothing can divert him from his path.

In reality, there are many different types of logic — more precisely, an infinite number. Therefore, we will not talk about "logic" in general, as is often done in some textbooks for lawyers, but about *mathematical logic*. After all, lawyers often have even less logic than politicians… And it is best to stay away from politics altogether.

Mathematical logic, too, is an ambiguous concept since there are infinitely many varieties of it. Here, we will discuss some of them, giving more weight to tradition than to common sense. Perhaps, paradoxically, that is where true common sense lies… Logical, isn't it?

Mathematical logic does not teach logical reasoning any more than any other branch of mathematics. This is because the "logicality" of reasoning in logic is defined by logic itself and can only be properly applied within logic itself. In real life, when we think logically, we usually use different types of logic and various methods of logical reasoning, freely mixing deduction with induction. Moreover, in practice, we often base our reasoning on contradictory premises, such as "Do not put off until tomorrow what you can do today" and "Haste makes waste." It is also common for an unwelcome logical conclusion to lead us to reconsider the initial premises (axioms).

Perhaps it is time to state one of the most important aspects of logic:

> *classical logic does not concern itself with meaning — neither common sense nor any other kind.*

The study of common sense, incidentally, falls under psychiatry. But in psychiatry, logic is more often harmful than helpful. However, some approaches, such as repertory grids[4], claim to have made progress in this area.

Of course, when distinguishing logic from meaning, we primarily refer to *classical logic* and the everyday understanding of common sense. In mathematics, there are no forbidden directions, so the study of meaning through logic, and vice versa, appears in various modern branches of logical science. (That last sentence turned out well, although I would not dare to define "logical science" even approximately.)

MEANING — or, if you prefer, SEMANTICS — is studied, for example, in MODEL THEORY. More generally, the term *semantics* is often replaced with *interpretation*. If we accept the philosophical view that interpretation (mapping!) of an object is the process of assigning meaning to it within a given framework, then the mathematical fields that can be brought to bear on the study of meaning in logic become vast and boundless!

In practical terms, theoretical programming is inevitably concerned with semantics. Moreover, in addition to basic semantics, it includes operational, denotational, procedural, and many other types of semantics.

Let us also briefly mention the pinnacle of this pursuit —

[4] Repertory grids are a psychological tool used to explore how individuals perceive and interpret their world. They were developed by George Kelly in the 1950s as part of his *Personal Construct Theory* (PCT). Repertory grids are designed to uncover the personal constructs (mental frameworks) that people use to make sense of their experiences and relationships. – DM

CATEGORY THEORY, which has taken semantics to the level of a highly formal and barely comprehensible syntax. Here, meaning is so meticulously structured and compartmentalised that it becomes nearly impossible for an ordinary person to grasp. This is reserved for the select few.

So, what exactly does logic concern itself with — at least in its most classical form? Logic concerns itself only with what it defines as its domain. (And it defines this with the utmost precision.) The essence of logic is to be strictly defined — to establish an *axiom system*. From that point on, logical inferences *must* be, to a large extent, automatic. However, reasoning *about* these inferences is a different matter altogether! Such reasoning falls outside the scope of logic itself. This is why it requires a rigorous mathematical foundation!

It might seem like mere verbal juggling. *NO!*

As an example of a logical (axiomatic) system, let us consider the well-known 15-puzzle. We begin by setting an initial arrangement of the numbered tiles (shuffling them). From this point, the game itself — akin to logical inference — consists of moving tiles into the empty space. A mechanical device could theoretically carry out these moves, while you watch patiently, rejoicing when the tiles eventually form the correct sequence from 1 to 15. However, nothing prevents you from guiding the device, using *common sense* to suggest optimal moves and speed up the process. Or, perhaps, you might even *prove* — using a branch of mathematics such as COMBINATORICS — that the given initial arrangement makes it impossible to achieve

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

The 15-puzzle is a sliding tile game on a 4 × 4 grid with 15 numbered tiles and one empty space. The goal is to arrange the tiles in order using sliding moves. Not all scrambled positions are solvable—solvability depends on the parity of permutations.

the desired final configuration at all!

No greater degree of common sense is found in that branch of logic known as BOOLEAN ALGEBRA. Here, logical operations are introduced and their properties are defined. As practice has shown, in some cases, the laws of Boolean algebra align with the *logic of* life, while in others, they do not. Due to this inconsistency, these logical laws cannot be regarded as universal principles applicable to real-world reasoning. Their knowledge and mechanical application may not only be helpful but also harmful — especially for psychologists and lawyers.

The situation is further complicated by the fact that, alongside the laws of Boolean algebra — which sometimes correspond to real-life reasoning and sometimes do not — there exist logical principles that some logicians outright refuse to acknowledge. This applies primarily to the *law of the excluded middle* and the *law of non-contradiction*. We will discuss these in greater detail later.

7 Operations On Statements

Mathematical logic is usually introduced through PROPOSITIONAL ALGEBRA, often recalling George Boole, the father of Lilian Voynich, who wrote *The Gadfly*. Boole himself, unburdened by mathematical erudition, attempted to create a mathematical system that could describe thought processes. Modern mathematical logic traces its origins to his "algebra."

Interestingly, many mathematicians do not consider this algebra to be true logic.

A STATEMENT (or PROPOSITION) is understood as a declarative sentence that can be either true or false. For example:

§ "The Volga flows into the Caspian Sea."
§ "The square of the hypotenuse is equal to the sum of the squares of the legs."
§ "Napoleon was born in Kolkata."

The first two statements are true, while the third is false.

Of course, real-life situations sometimes complicate things. For instance, the statement about the Volga is true if one knows this fact from geography. I once had to explain to an American businessman that far from the United States, there exists a large river called the Volga… Similarly, not everyone can confidently confirm the Pythagorean theorem.

However, let us agree not to consider uninformed individuals. To avoid getting lost in minor, irrelevant details, we will define statements as declarative sentences whose truth value can be determined by a "higher intelligence".

However, the problem does not end there. The declarative sentence "I am lying" is not a statement because if it is true (meaning I really am lying), then I am not lying but telling the truth! And vice versa… This is an example of a LOGICAL PARADOX.

Logical paradoxes are not considered statements. Similarly, questions, exclamations, and definitions (i.e., non-declarative sentences) are not statements either. It is meaningless to discuss the truth or falsity of definitions. A definition is simply an agreement on naming. For example, "Let us call this music an anthem." And that's it!

To avoid repeatedly writing "TRUE" and "FALSE", their initial letters are often used. Even more commonly, we simply use "1" and "0".

Now, let us return to the most essential point. Propositional logic does not deal with (or even concern itself with) the *meaning* of statements. In this sense, logic can be considered *nonsense*! One of the classical logicians compared propositional algebra to an X-ray, which, when scanning a statement, leaves only its truth value for mathematical examination.

In propositional algebra, two or three operations are usually sufficient, though more are typically studied. The operation *disjunction* is also called "LOGICAL OR". When two statements are connected by disjunction, the resulting compound statement is TRUE if at least one of the original statements is TRUE. That is, it is a "NON-EXCLUSIVE OR".

For example, "We love beer or we love ice cream" is a true compound statement as long as *at least* one of the simple statements is true — or perhaps even both. The idea of a living being that loves neither beer nor ice cream is beyond imagination!

The CONJUNCTION operation is also called "LOGICAL AND". A compound statement formed using conjunction is true *only if* both of its component statements are true.

The negation operation — "LOGICAL NOT" — converts a true statement into a false one and vice versa.

Perhaps the most intriguing operation is implication, or "LOGICAL IF ... THEN". For example, "If Napoleon was born in Kolkata, then gas contracts when heated." Interestingly, this is a true statement! There is no reason to consider it false.

The only situation in which an implication is false is when the premise (the "if" part) is true, but the consequence (the "then" part) is false.

Even more puzzling, from the perspective of common sense, is that implication is sometimes (not entirely correctly, for other reasons) called an operation of *logical entailment*. However, our example demonstrates that statements do not necessarily follow from one another logically, and in fact, they may have *no* logical connection at all.

It is important to remember that implication, like other logical operations, considers only the truth values of its components.

> "If the Volga flows into the Caspian Sea, then $2 + 2 = 4$" is a true statement.
> "If the Volga flows into the Caspian Sea, then $2 + 2 = 5$" is a false statement.

Although, from the standpoint of common reasoning, both of these "logical arguments" are equally meaningless.

There is also *logical equivalence*, or "IF AND ONLY IF" (incidentally, using an "American trick," this can be abbreviated as "IFF"). A compound statement formed using equivalence is true if both of its component statements are *either true or false at the same time*.

Another important operation is the SHEFFER STROKE[5], or "NAND" (LOGICAL "AND-NOT"). The result of this operation is equivalent to *applying conjunction followed by negation*. That

[5] The Sheffer stroke, also known as the NAND operation, is named after Henry Maurice Sheffer, who introduced it in 1913.

is, the resulting statement is false *only* if *both* of its components are true. The Sheffer stroke is remarkable because it is functionally complete — any logical expression can be represented *solely* using this operation, applied a sufficient number of times.

When logic was first applied to designing logical circuits, such as processor components, an analogy was initially drawn with relay circuits. The *disjunction* ("OR") operation corresponds to parallel relay contacts, the conjunction ("AND") operation to series-connected contacts, and the negation ("NOT") operation is modeled by a normally closed relay contact—meaning it opens when the relay is triggered. Naturally, all this was later implemented in semiconductor-based modular designs.

At that time, it was enough to produce NAND gate modules, which could be used to implement *any* logical circuit. (Although the processor itself was the size of a wardrobe — but not because of logic!)

8 Propositional Algebra

In this algebra, the objects are the propositions we have already discussed. We have also covered operations on propositions. What remains is to examine their properties or laws to formally define the algebra.

If we consider only the first three logical operations — disjunction, conjunction, and negation — then propositional algebra is analogous to set algebra. The analogue of disjunction (OR) is union, conjunction (AND) corresponds to intersection, and negation (NOT) is complementation. These analogies can be used to explain the meaning of logical operations — this is the set-theoretic interpretation, which is quite natural. However, we will focus on a formal approach.

We have also mentioned implication, equivalence, and the Sheffer stroke, but we did not seek set-theoretic analogues for them. However, these operations can be expressed in terms of

the first three.

For example, implication can be rewritten as the disjunction of the negation of the first proposition with the second. In formal logic, the following statements are equivalent:

> "IF the weather is good, THEN we go swimming."
> "It is NOT TRUE that the weather is good OR we go swimming."

The only case in which both statements are false is when the first proposition is true (the weather is good) and the second is false (we do not go swimming).

For *equivalence*, the transformation is longer but essentially follows the definition. For example, consider the (somewhat unusual) statement:

> "The weather is good IF AND ONLY IF we go swimming."

This is equivalent to the statement:

> "The weather is good AND we go swimming OR the weather is NOT good AND we do NOT go swimming."

Equivalence can also be expressed in terms of conjunction of two implications:

> "IF the weather is good, THEN we go swimming AND IF we go swimming, THEN the weather is good.

For the Sheffer stroke, applied to the same propositions, we can rewrite it as:

> "It is NOT TRUE that the weather is good AND we go swimming."

According to De Morgan's law, this is equivalent to:

> "The weather is NOT good OR we do NOT go swimming."

In propositional algebra, there are three fundamental laws: *commutative*, *associative*, and *distributive*, which are analogous to the corresponding laws in set theory.

To illustrate the commutative law, let us use an example from Kleene's book *Mathematical Logic*[6]:

> "Mary got married AND had a child."

From a logical point of view, this is equivalent to:

> "Mary had a child AND got married."

Here, two aspects are highlighted:

Syntactically, the commutative law allows us to swap statements. Semantically, the swap does not align with societal norms—order matters in social contexts. This demonstrates that mathematical logic does not (and cannot) account for many real-life nuances. The associative law states that the order in which we group conjunctions and disjunctions does not affect their truth value. For example:

> "The weather is good AND we go swimming AND we catch a cold."
> "The weather is good OR we go swimming OR we catch a cold."

[6] S.C. Kleene, *Mathematical Logic*, John Wiley & Sons, 1967.

Since mathematical notation often uses parentheses to indicate the order of operations, the associative law is also called the LAW OF REMOVING PARENTHESES.

Distributive Law: Let us consider an example of a more "exotic" case.

> "The weather is good OR we go swimming AND we catch a cold" is logically equivalent to:
> "The weather is good AND we go swimming OR the weather is good AND we catch a cold."

We will not list all possible laws of propositional logic, as they closely resemble the laws of set algebra. However, it is important to note that instead of using the word *equality*, we have been using *equivalence*.

Two compound statements are *equivalent* if they have the same TRUTH TABLES. In these tables, the initial columns correspond to the elementary statements, while the final column corresponds to the compound statement. The initial columns list all possible truth value combinations of the elementary statements, and the final column reflects the truth value of the compound statement for each combination.

Each row represents a separate combination of truth values.

For the last example, the truth tables for both the left and

right sides of the distributive law are identical:

Good Weather	We Swim	We Get a Cold	Result
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

Regarding mathematical logic, as with set theory, there are people who disagree with some of its laws. The primary objections are again directed at the laws of *excluded middle* and *non-contradiction*. That is, for constructivists (intuitionists), the completion of seemingly obvious truth tables is not so obvious!

Constructivists argue that the complex statement "Fermat's theorem is true OR Fermat's theorem is NOT true" cannot be true, if only because accepting its truth would make the question "So is it true or not?!" permanently unsolvable. More humanistic logicians use examples such as "A person is almost bald OR it is NOT TRUE that a person is almost bald." They claim that determining the truth of this complex statement is not only impossible but also simply tactless.

Propositional logic, and propositional algebra in particular, as mentioned earlier, flourished during the early days of computing. One of the most important algebraic transforma-

tions is the *minimisation of complex statements*. That is, many methods were developed to derive from an initial statement an equivalent one with the smallest possible number of logical operations. A technical device can then be built according to such a statement with maximum simplicity. All interested parties benefit from the results obtained through science.

9 Predicate Logic

The term "predicate" translates from English as *verb*. However, saying *logic of predicates* would be disrespectful to oneself. Formally, a *predicate* is a function whose arguments can be *any objects* from a certain set, and the function's values are TRUE or FALSE. A predicate can be viewed as an extension of the concept of a statement.

Example. Instead of three separate statements:

> "Masha loves porridge."
> "Dasha loves porridge."
> "Sasha loves porridge.",

one can write a single predicate:

> "*X* loves porridge."

and agree that *X* can be replaced with either Masha, Dasha, or Sasha.

Substituting a specific child's name for *X* turns the predicate into an ordinary proposition.

For predicates, the logical operations discussed earlier still apply and retain their meaning. For example:

> "IF Masha loves porridge, THEN Sasha loves porridge."

There are also two new operations that are specific to predicate logic. These are known as *quantifiers*, corresponding to the phrases "FOR ALL" (*universal quantifier*) and "THERE EXISTS" (*existential quantifier*). We agreed not to write formulas, but it is worth mentioning the symbols used, as they are quite distinctive. The universal quantifier comes from the English word "All" and is represented by an inverted "A" (∀). The existential quantifier comes from "Exist" and is represented by a rotated "E" (∃), as flipping it upside down would not work.

When a quantifier is applied to a predicate, the result is a proposition, which can be either true or false:

> "EVERYONE loves porridge"
> "SOME people love porridge"

The predicate in this example, before quantification, is a *one-place predicate* (a function with one argument). However, predicates can have more than one argument. This is easy to illustrate if we consider that children might love not just porridge but other things as well.

For example, "*X* loves *Y*" is a *two-place predicate* (binary predicate).

> "EVERYONE loves *Y*" is a one-place predicate (after quantification over *X*).
> "EVERYONE loves SOMEONE" results in a zero-place predicate, which is simply a proposition.

It is interesting to examine how quantifiers behave in the presence of negation. Consider the negation of the predicate "EVERYONE loves porridge":

> "It is NOT TRUE that EVERYONE loves porridge."

By De Morgan's law, this is equivalent to the statement "SOME do NOT love porridge." In other words, the negation moves past the quantifier, switching it to its opposite.

Now, let us make one of the most important statements:

> **Among the formalized languages of mathematics, predicate logic is the closest to natural language.**

This is why artificial intelligence research tends to favor its use. Compared to natural language, it is restricted in many ways, yet no better alternative has been found in over a century — except for so-called SYNTACTIC SUGAR,"[7] where natural language words replace symbolic notation (as we are attempting to do here).

In well-formalized systems, the language is often further restricted to facilitate implementation on computers. A key example of this is the PROLOG programming language — short for *Programming in Logic*.

[7] The term "syntactic sugar" was introduced by Peter J. Landin in 1964 to describe syntax improvements that make code or logical expressions more readable without altering their functionality. In predicate logic, syntactic sugar includes replacing formal symbols with natural language phrases, enhancing readability while preserving formal structure.

Thus, predicate logic inherits the flaws of propositional logic, which cause an inherent mismatch between natural and logical languages.

Predicate logic can describe a great deal, though not everything. Even within its limitations, one often has to resort to tricks and techniques that resemble craftsmanship or art rather than strict formalism — though explanations, in the end, are usually rigorously structured.

Here are some "classic examples."

If we want to express in predicate logic that "All students are excellent," the recommended construction is:

> "For all x, IF x is a student, THEN x is excellent."

However, if we want to say, "Some students are excellent," we should write:

> "For some x, x is a student AND x is excellent."

The conditional "IF ... THEN..." does not work in this case. The reason is that if even one non-student is included, the statement becomes true regardless of whether there are any excellent students.

Another example is the statement: "Dogs and cats are not allowed inside." The construction:

> "For all x, IF x is a dog AND x is a cat, THEN x is not allowed inside."

Clearly, no such x exists that is both a dog and a cat simultaneously! The same applies to any y. Therefore, the correct formulation should be:

> "For all x, IF x is a dog OR x is a cat, THEN x is not allowed inside."

The list of such logical tricks and workarounds could go on indefinitely. However, the key is to recognize that even if one were to invent an alternative logic that avoids these issues, it would likely introduce new, potentially greater problems.

And we have not even touched on the complexities inherent in natural languages — such as synonyms, homonyms, metaphors, hyperboles, and so on. Idioms alone are an entirely different challenge! So, any "reconciliation" between logical and natural language remains unlikely, even in the most distant future — perhaps not even when the Earth itself begins to cool…

10 Axiomatic Theories

If Boolean algebra provides a model of logical thinking, it is a very specific one. Strict logical reasoning does not even come close to such a model. The deductive approach allows for the creation of (absolutely) rigorous logical (that is, absolutely logical) systems. This should not be confused with Sherlock Holmes' "deductive method." Holmes, or rather Conan Doyle, clearly had problems with logic, as he confused deduction with induction…

The DEDUCTIVE approach, also called the AXIOMATIC approach, moves from the general to the particular, from axioms (postulates) to theorems (consequences). An axiomatic theory is strictly defined if its language, axioms, and rules of inference are precisely formulated. If even one of these three components is weakened, the strictness of the theory collapses!

The famous (and one of the first) axiomatic theories —

Euclidean geometry — managed to maintain strictness only in one component: its postulates. However, the language in which geometric proofs are conducted, even after millennia, and the strictness of the proofs themselves, do not withstand scrutiny. It is nothing more than an ambiguous, metaphorical-allegorical language and plausible reasoning. This is why cases of "proven" theorems being refuted are not uncommon. In fact, almost all of mathematics, except for a comparatively tiny branch in the logic of axiomatic systems, rests (or, better put, is founded) on plausible reasoning and the integrity of those proving the theorems.

Thus, an exemplary, impeccably rigorous theory is formulated in the language of predicate formulas. (We have vowed not to use formulas here, so we must rely on our own honesty.)

Some formulas are declared to be AXIOMS. In life, too, we do not declare every phrase we can come up with in a given situation to be a law (axiom), but only certain ones that we decide to consider as such… Laws (axioms) are a matter of belief, and sometimes of practicality. They are *unprovable*! If they were provable, they would already be theorems!

The existence of God is unprovable! Otherwise, it would be a theorem. And from which, pray, more fundamental concepts would one derive such a "theorem"?…

The law of universal gravitation is unprovable. We simply chose to believe in it because we got tired of conducting experiments on falling bodies, waiting for something original to

happen to them.

Inferences in a theory must also be formalized, as everyone has encountered people in life who "convincingly" prove complete nonsense.

By the way, the most famous rule of inference in mathematical logic — *modus ponens* — is depressingly obvious and even primitive. It can be illustrated as follows:

Suppose a system contains the statements

> "IF the weather is good, THEN we go for a walk"
> AND
> "The weather is good"

Then, according to *modus ponens*, the conclusion follows:

> "We go for a walk."

Despite its primitiveness, this rule of inference has a crucial advantage. It is obvious to everyone. Nothing could be more obvious! And IF the system also contains the statement:

> "IF we go for a walk, THEN we will definitely get lost."

THEN, given the previously derived statement

> "We go for a walk,"

we obtain:

> "We will definitely get lost."

See how far one can go with small, obvious steps!

There are many other rules of inference, but they all share one essential property—obviousness. This obviousness allows these rules to be used in an entirely formal manner, meaning the result is computed. Such symbolic computations are called FORMAL CALCULI.

There is another approach to axiomatics that focuses primarily on rules of inference. These systems (for some reason) are called NATURAL DEDUCTION SYSTEMS, implying that they encompass the fundamental natural rules of logical reasoning.

Logicians played among themselves until the approach to constructing axiomatic systems known as the PRINCIPLE OF RESOLUTION (or RESOLUTION METHOD) was formulated. This greatly contributed to bringing logic to the broader public.

On one hand, it spurred work on using computers for logical inference and artificial intelligence research in particular. On the other hand, it led to the creation of the PROLOG programming language.

This is a completely different kind of programming compared to traditional *procedural* programming. Prolog is a *declarative* programing language. It is not even programming in the usual sense of the word, since the programmer does not write an algorithm to solve a problem. Instead, they describe the logical dependencies of the "world" in which the problem exists. Based on this logical description, the system (machine) constructs the algorithm itself in the process of searching for

a solution!

It may seem that axiomatic systems are complex, but in reality, anyone can invent as many as they like. You have probably never done anything simpler.

For example, let us declare that any "word" consisting of the letter ℜ belongs to our language. Let us declare the letter ℜ as an axiom. Our inference rule will be: double the letter ℜ.

This means we have immediately created a theory in which any sequence (word) composed of the letter ℜ is derivable:

ℜ, ℜℜ, ℜℜℜ … ℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜℜ …

All would be well, except that such a strictly defined theory offers little to its creator beyond the joy of creation. This raises the question of purpose and meaning — that is, *semantics*. Here, logicians have taken a very (!) interesting position.

Since logic is not concerned with the *meaning* of statements but only with their *truth*, truth itself was declared to be their meaning. Think about it: the meaning of the statement "Gas expands when heated" is not that it reflects a physical law but simply that it is *true*. Consequently, the statement "Nikita Mikhalkov is a film director"[8] has exactly the *same meaning* (!) — that it is also true.

And it goes even further. The predicate language is a significant extension of the propositional language, and even in simple cases like "*X* likes porridge", it is not always possible to

[8] Nikita Mikhalkov is a renowned Russian filmmaker, actor, and public figure. He is best known for directing critically acclaimed films such as Burnt by the Sun (1994), which won the Academy Award for Best Foreign Language Film. He is a prominent cultural figure in Russia and has been involved in both the arts and politics. – DM

enumerate all possibilities. This becomes even more problematic when dealing with infinities. To address semantic issues in such cases, MODEL THEORY is employed. However, in the end, this theory also reduces "meaning" to the TRUE-FALSE dichotomy.

Returning to axiomatic theories, it should be noted that in mathematics, only those theories that allow the derivation of true formulas — and prevent the derivation of false ones — have *practical meaning*.

One false formula *completely annihilates* any axiomatic theory.

Our theory, built from the letter ℜ, is not tied to the concept of truth. Therefore, it is meaningless — akin to the random rearrangement of children's building blocks.

Building blocks are straightforward. However, the challenges of axiomatic theories do not end there. Perhaps the most fundamental discovery in this field is GÖDEL'S INCOMPLETENESS THEOREM. It turns out that in any sufficiently complex axiomatic system (more complex than building blocks but as simple as arithmetic), there exist formulas that can neither be proven nor disproven. Could this be why not all school problems have solutions?

So, when constructing your own axiomatic theories, remember that they should possess some useful properties. And creating such theories is not so easy. However, anyone can create their own mathematics!

A well-known mathematician once said: “The advantage of axiomatization is the advantage of theft over honest labour.”

11 Theory Of Algorithms

The THEORY OF ALGORITHMS does not teach how to construct algorithms. It deals with a more fundamental question. The primary task of classical algorithm theory is to answer the question:

> "Is it (in principle) possible to construct an algorithm for problems of this type?"

Or, in more formal terms:

> "Are problems of this type algorithmically decidable?"

This is important for two reasons. First, not all problems can have algorithms for their solution. Second, to rigorously prove that constructing an algorithm is *impossible*, one needs a strict (formal) definition of the algorithm itself. However, the concept of an ALGORITHM is a fundamental, undefined notion. When it comes to algorithms, we are like a dog: we

understand, but we cannot explain. If you ever come across a "definition" of an algorithm, rest assured that every word in it is just an analogy.

A non-trivial way out of this impasse was found. The concept of an algorithm was replaced by strictly formalised mathematical models. Among the most well-known are RECURSIVE FUNCTIONS, TURING MACHINES, and NORMAL MARKOV ALGORITHMS.

These mathematical models serve as concretisations of the concept of an algorithm. Long-standing practice supports the so-called CHURCH'S THESIS, which can be restated as follows:

> *For any algorithmically decidable problem, a recursive function (Turing machine, Markov normal algorithm) can be constructed. Conversely, for problems where none of these concretisations can be built, no algorithmic solution exists.*

Recursive Functions: Recursive functions are based on the idea that the input data and possible results of solving any problem can be assigned numerical values. For this, the set of natural numbers (positive integers starting from zero) is sufficient. The fundamental functions, whose computability is unquestionable, are then declared as basic.

Zero Function: This function returns zero for any input. Even a child can implement this function. One could even train a parrot to scream "Zero!" in response to any query about its value.

Successor Function: This function provides the next value relative to its argument. For five, it returns six; for a million, it returns a million and one. One might say that this is simply adding one. However, at this stage, we have not yet defined addition!

Projection Function: This is an amusing function, even for a first-grader. It is named after the position of an argument and returns the value of the specified argument. For example, if we have several arguments — let's say Ivanov, Petrov, and Sidorov — numbered as 22, 13, and 49 respectively, then the function that selects the third argument will return 49.

These three basic functions serve as the building blocks for constructing more complex functions using three operators: COMPOSITION, PRIMITIVE RECURSION, and MINIMISATION.

The COMPOSITION OPERATOR, well known from school, allows substituting a function in place of an argument — like a nested structure: "The needle is in the egg, the egg is in the chest…"

Describing the **Primitive Recursion Operator** takes longer, but with some effort, it can be understood. This operator enables the construction of a new function from two functions: one with one argument fewer and another with one argument more.

The value of the newly created function for an argument equal to zero is set equal to the function that lacks this argument. The value of the newly created function for all other

(non-zero) values of the selected argument is defined through another function, which depends on the same arguments except for the selected one, on the *previous* value of the selected argument, and on the newly created function applied to the previous value of the selected argument.

At this point, one cannot help but regret the absence of formulas.

However, in reality, we have encountered such functions since school, though, like a certain literary character, we only realise later in life that we have been speaking prose all along.

To illustrate, let us construct the well-known two-variable function of multiplication, $x \times y$, using recursion (taking y as the selected variable).

The function for multiplying x by zero can be expressed using the zero function of x, which ensures the desired value—zero.

The function for multiplying x by y (where y is not 0 can be expressed through the function of adding x to the value of the function for multiplying x by the previous value of y … In other words, we have expressed multiplication in terms of addition.

Two remarks are necessary here. First, we assume that by this point, the recursive nature of the addition function used here has been proven. Second, when multiplying x by y, we have at our disposal a function of three arguments. However, by applying a projection function, we can eliminate the middle

argument, as it is unnecessary in this case. After all, we are permitted to select from what is possible!

The final operator is the **Minimisation Operator**. Its necessity can be explained simply by the fact that recursive functions, which are designed to solve any algorithmically decidable problems, themselves only operate with positive integers. This limitation prevents even a simple arithmetic problem such as 5 – 8 = ? from being solved — recursive functions do not accommodate negative numbers.

Of course, these elementary problems can be handled in a childlike way. One could agree that 1000 serves as a "shifted" zero and that subtraction is effectively addition with a "negative" number. Then 1005 + 992 = 1997. (Adjusting to the scale: 1997 – 1000 = 997).

Since the zero shift was accounted for twice, the final result (in a system where zero has been shifted once by 1000) will be 997.

However, this childlike workaround merely demonstrates that even in school, one can manage without negative numbers — just as ancient Rome did. However, unlike the Romans, we are in dire need of zero.

It is the minimisation operator that tracks the value of the chosen argument at which the function it observes first reaches zero. This value of the chosen argument is the value of the minimisation operator. For example, for the function $x - y$ with $x = 5$, the value of the minimisation operator is also 5,

since, as we increase y from zero, the function reaches zero precisely when $y = 5$.

The basic functions, along with those that can be constructed from them using the operators of composition, primitive recursion, and minimisation, form the set of *partial recursive functions*. The set of partial recursive functions coincides with the set of all algorithmically solvable problems (the set of all computable functions). Incidentally, it is the minimisation operator that is responsible for the term "partial", as it introduces functions that are not defined everywhere (partial functions) into the set.

Turing was not an automobile engineer. The TURING MACHINE does not require an internal combustion engine, as everything in it moves solely by the power of thought. It is a mathematical model. It may, in some ways, resemble a machine, but no more than a tape recorder resembles an automobile. In a Turing machine, the tape (divided into cells) remains stationary, while the read-write head moves along it. Worse still, the head moves in jerks, from one cell to another. Symbols are recorded in these cells, and to ensure there are no empty cells, a special blank symbol is written in them.

A Turing machine has a control unit with a memory of "states" and operates according to a given program (algorithm). The program consists of commands. Each command involves the following steps: The machine reads the symbol in the cell where the head is positioned (while being in some state [initially in the start state]), writes a symbol into this cell (which

may be the same as the one read), changes its state (which may remain the same), and moves one step left or right (or may stay in place).

The machine moves along the tape until it reaches a special state known as the final state. This marks the end of the machine's (or algorithm's) execution. The result (solution) remains on the tape.

Example. Let us construct a machine that erases the last one in a continuous sequence of ones. Since the number of ones in the sequence is arbitrary and unknown, we define the last one as the one that appears *before* the blank symbol. This is the core idea of the solution. The rest is a matter of implementation. The program consists of four commands:

- The machine reads a blank symbol while in the initial state, writes a blank symbol, and moves one step to the right. (This means the machine is positioned *before* the sequence of ones.)
- The machine reads a one while in the initial state, writes a one, moves one step to the right, and remains in the same state. (This means the machine is "traversing" the sequence of ones.)
- The machine reads a blank symbol while in the initial state, writes a blank symbol, moves one step to the left, and transitions to the second state. (This means the last one has been found.)
- The machine reads a one while in the second state, writes a blank symbol (erasing the one), remains in place, and

transitions to the final state. (The task is complete.)

Despite the apparent simplicity of this construction, *a Turing machine can be built for any algorithmically solvable problem!* And since the machine is constructed entirely in one's mind, concerns about its "technical efficiency" are irrelevant. The only question that matters is whether the machine will eventually reach the final state — even if only after an *imaginary* million years. If so, the problem is solvable!

It would not be an exaggeration to say that normal algorithms were created by A.A. Markov, a corresponding member of the USSR Academy of Sciences from Moscow. To ensure uniformity, he, as the rightful author, called them *algorifms* (алгорифмы) instead of *algorithms*, since the word has both Arabic and Greek origins — just like the word *arithmetic*.

The essence of normal algorithms lies in *forced substitution*, where the order of replacements is strictly defined. In normal algorithms, the algorithm itself is specified by a *normal substitution scheme* — a sequence of rules dictating "what to replace with what." This is best illustrated with an example involving word substitutions, especially since Markov himself referred to any sequence of letters, even those not found in any dictionary, as "words."

For instance, given two substitution rules: replacing "mo" with "ele" and "use" with "phant", one can transform "mouse" into "elephant".

The mechanism of normal algorithms is so simple that it re-

sembles a children's game rather than mathematics. However, it is actually a very powerful mechanism, as it can be used to express the solution to any algorithmically solvable problem. Once again, it is important to note that this should not be taken as a suggestion to solve every problem through substitutions (although the remarkable programming language REFAL[9] is based on these principles). It simply means that any algorithmically solvable problem *can* be represented as such a system of substitutions. And if it *cannot* — and you can prove this — then the problem has no algorithmic solution at all.

[9] REFAL is a functional programming language that uses recursive functions and substitutions to solve problems. It was developed in the Soviet Union in the 1960s by Valentin Turchin, a prominent computer scientist. It is based on the principles of normal algorithms and Markov algorithms, which are formal systems for string manipulation and computation. REFAL is designed to work with symbolic computations and is particularly well-suited for tasks involving pattern matching, symbolic transformations, and recursive problem-solving. – DM

12 Formal Grammars

FORMAL GRAMMARS are a well-developed mathematical framework that, aside from studying abstract concepts, allows for the mathematically rigorous creation of programming languages and the development of compilers for these languages.

There is an insurmountable gap between natural and formal languages, so any similarity in terminology should be considered coincidental… Moreover, within the vast and intricate LANGUAGE OF MATHEMATICS, the study of formal grammars and languages is primarily focused on compiler construction.

A formal language can be defined as a specific set of strings. A string is a sequence of symbols. In this context, any computer program is also treated as a string. Spaces in a program are special symbols, for which the longest key on the keyboard is reserved.

However, not every random sequence of characters producible by a keyboard qualifies as a valid word in a given language — only lexically and syntactically correct programs do. A program that is grammatically flawless may still be useless, meaningless, or even harmful. However, ensuring a program's correctness in execution is beyond the responsibility of formal grammar and the compiler. (Mathematics, after all, generally does not concern itself with meaning.)

In formal grammars and languages, meaning is also outside the scope of mathematics. There is a specific branch of theoretical programming where a program's intended function is described in a formal language, usually predicate logic or its dialects. Based on this description, a specialized system synthesizes the program. However, this is a separate topic entirely—especially considering that people tend to make more mistakes when describing what a program should do than when writing the program directly.

To define a grammar, it is necessary to specify the sets of TERMINAL and NON-TERMINAL symbols. Terminal symbols are those used in the language, while non-terminal (intermediate) symbols are used in the generation (derivation) of words in the language. Words are created according to GRAMMATICAL RULES. Each word, from a programmer's perspective, is essentially a program written exclusively in terminal symbols.

Next, GRAMMATICAL RULES are established. These are quite similar to the substitutions used in Markov algorithms. However, unlike the latter, the order in which grammatical

rules are applied is arbitrary. Applying a rule consists of replacing a sequence of symbols in a transformable string — one that matches the left-hand side of a given rule — with the right-hand side (a sequence of symbols) specified by that rule.

For purely aesthetic reasons, let us introduce another elegant term — SENTENTIAL FORM. In constructing programs within formal grammars, one always starts from a single initial non-terminal symbol. Let us denote this symbol as "program". A substitution then occurs according to one of the grammatical rules, replacing this symbol with the corresponding right-hand side, which may include a sequence of both terminal and non-terminal symbols. This process is called *immediate derivation*.

Any of the resulting non-terminal symbols can be replaced with a sequence of symbols according to an applicable grammatical rule. That is, the initial non-terminal symbol "program" successively transforms into an increasingly long sequence of symbols until, at last, only terminal symbols remain. At this point, a word of the given language has been obtained — ironically called a "sentence." All sequences of symbols that appear in the course of immediate derivations, existing between the initial non-terminal symbol and the final sentence, are referred to as SENTENTIAL FORMS. And all that remains for us is to appreciate the fact that English is not our native language.

A compiler, upon receiving a program, parses the given sentence by applying the grammatical rules in reverse (moving from the right-hand side of the rule to the left) until it reaches

the start symbol “Program”.

There are usually a vast number of ways to both generate and parse a sentence. If the parsing process fails, alternative options must be explored. A word will be considered *not* belonging to a given language (grammar) if none of the parsing attempts succeed. Since an exhaustive search of possibilities is generally impractical, grammars are designed with *useful* properties rather than being arbitrarily defined. Parsing (recognition) methods leverage these beneficial properties to minimise or entirely eliminate unnecessary backtracking.

There is a rather crude but still useful classification of grammars introduced by Noam Chomsky. He divides them into three types, excluding the zeroth type. The zeroth type consists of grammars with completely unrestricted grammatical rules. Since there are no constraints, anything can be found within them, making analysis impossible. Or rather, one might consider the analysis complete with the conclusion:

> *Anything and everything—desirable or undesirable—can exist here.*

Dealing with such grammars is therefore pointless. Even for the insane.

Grammars of the third type are called AUTOMATA BASED or REGULAR grammars. This is because they are generated and recognised by *finite automata* (this mathematical model is associated not with Kalashnikov but with the names of mathematical logicians such as Mealy, Moore, Trachtenbrot, etc.)

and by *regular expressions* (similar to a regular army, where expressions are constructed according to simple rules and easily recognised — this too is a mathematical model).

Automata based grammars are typically used at the lexical level. A *lexeme*, in the usual sense, is a unit of vocabulary. However, from the compiler's perspective, it is a "symbol", since the entire program is considered a word. For instance, the number 345.08 may be recognised as a single symbol — a floating-point number.

Lexical analysis in a compiler precedes syntactic analysis. Famous UNIX tools such as `lex` and `yacc` help automate the creation of lexical and syntactic analyzers for compilers.

We have strayed too far into programming. Programming is also mathematics — discrete mathematics — but a different kind. And that is another story.

And from programming, one usually does not return so easily…

13 Computational Complexity

Not only are there algorithmically undecidable problems, but they are infinitely more numerous than algorithmically decidable ones. From a practical perspective, it does not help us much if the solution to a decidable problem can only be obtained in a million years. The calculations simply cannot be completed earlier… These are *hard-to-solve* problems. For us, mere mortals, such problems are no different from algorithmically undecidable ones.

The situation with these problems is even more unclear than with algorithmic decidability. So far, the only "solid" argument for our helplessness in the face of a hard problem is that no one has yet found an easy solution for it — even American politicians.

The abnormality of this situation becomes even more pronounced when we consider that theorists only deal with *finite*

discrete problems — problems of either *optimisation* or *decision-making* (YES/NO classification) — each of which, in theory, can at least be solved by *brute force*, if no better method exists. But that does not make things any easier if we recall the legend of the inventor of chess, who asked for one grain of wheat for the first square of the chessboard, two for the second, and so on — until the 64th square required **2 to the power of 64 grains**! This far exceeds the world's grain resources.

One book on computational complexity began with a quote from the Ukrainian philosopher Hryhorii Skovoroda:

> "Thank you, Lord, for making all that is necessary easy, and all that is difficult unnecessary."

However, a more fitting thought from Skovoroda might be the epitaph on his grave:

> "The world tried to catch me, but did not succeed."

And a more specific reflection from modern mathematicians:

> "Complexity is becoming the problem of the century."

Due to the overwhelming number of irrelevant details in different computational methods (approaches, paradigms), the characteristics that should be considered in defining computational complexity are too numerous and contradictory. However, in the end, most agree that everything can be reduced to *computation time* (number of elementary steps) and

memory usage. Moreover, many consider TURING MACHINES to be the best models for this purpose.

The problem has quickly been reduced to two key terms: POLYNOMIAL and NON-DETERMINISTIC.

Consider a large bag of stones, and suppose we need to find the largest one. We take out the stones one by one, comparing each with the largest found so far. By the time the bag is empty, we will be left holding the largest stone. If we double the number of stones, the time required to solve the problem also roughly doubles. If the bag contains n stones, the difficulty of solving the problem is proportional to n. Such a problem is said to be solvable in *polynomial time*.

If instead, we need to find the largest edge in a complete weighted graph, the problem is still of polynomial complexity, but its difficulty increases following the formula for a complete graph: as the number of vertices grows, the number of edges grows proportionally to n square.

However, there are problems (even for graphs) for which *no* simple polynomial-time solutions have been found. The second classic example of such problems (we will save the first one for later) is the TRAVELLING SALESMAN PROBLEM (TSP):

> *What is the shortest cycle in a weighted graph that visits each vertex exactly once?*

This problem has only *hard solutions*, with complexity growing *exponentially* (like the number of grains on a chess-

board).

However, verifying a solution to such a problem *can* be done in polynomial time. The main difficulty lies in finding the correct solution path. The NON-DETERMINISTIC TURING MACHINE (NTM) helps eliminate this problem of choosing the right trajectory. One can imagine an *arbitrarily* large number of deterministic Turing machines, each attempting to reach the solution along a different possible path. For those struggling to visualise this, one can simply imagine an oracle (often referred to as "the oracle") that reveals the correct path, avoiding the need to explore countless unknown routes. The final effect remains the same.

Thus, the set of *hard-to-solve problems* (NP problems) consists of problems that can be solved by a *non-deterministic machine* in *polynomial time*. Mathematicians have expressed the problem of computational complexity in a simple equation, which I will include for its brevity and ease of notation:

$$P = NP?$$

Mathematicians ask:

> *Do the set of problems solvable in polynomial time and the set of NP problems coincide? Or are we simply lacking the insight to find polynomial-time solutions?*

Regardless, problems for which *no simple* (polynomial-time) solutions have been found *do* exist. And as time passes, mathematicians seem ever more convinced (though without

proof) of this fact. Moreover, they have compiled a collection of typical NP-HARD PROBLEMS, now numbering at least a thousand.

Furthermore, it is claimed that some NP-hard problems reduce to others, leading to the classification of NP-COMPLETE PROBLEMS. A radical claim follows:

> *If a polynomial-time solution is found for even one* NP-COMPLETE PROBLEM, *then polynomial-time solutions will exist for all* NP-COMPLETE PROBLEMS.

In that case, the $P = NP$ conjecture would be proven true, and this formulation of the computational complexity problem would be resolved!

Afterword

The very first NP-COMPLETE PROBLEM was the problem of finding *cliques* in a graph, that is, complete subgraphs of a given graph with a specific number of vertices…

However, in the mid-1970s, so-called MAGÓ ALGORITHMS were published, which eliminated not only problems like the "Eight Queens Puzzle" (previously a standard testing ground for heuristic algorithms in "artificial intelligence) from the category of brute-force problems but also made it possible to find graph cliques through straightforward manipulations at the level of *propositional algebra* (transforming expressions into DISJUNCTIVE NORMAL FORM (DNF). These manipulations do not exceed polynomial complexity.

Admitting one's own ignorance and incompetence is hardly a source of joy, but for me, the problem of hard-to-solve problems exists in a somewhat distorted form.